CMP

WHEN BRAINS COLLIDE!

Dr. Dobb's Portal

Site Arcl'

ABOUT US | CONTACT | ADVERTISE | SUBSCRIBE | SOURCE CODE | CURRENT PRINT ISSUE | FORUMS | NEWSLETTERS | RESOURCES | BLOGS | PODCASTS | CA

Architecture & Design

June 01, 2002

Cutting Edge Server Farms

The blade server debate

* Email   * Print
* Discuss   * Reprint

add to:
* Del.icio.us   * Slashdot
* Digg   * Y! MyWeb
* Google   * Blink
* Spurl   * Furl

Lamont Wood

Server blades have caught the attention of administrators operating high-density server farms, but they may not be right for everyone. As with any new technology, there are several trade-offs to consider.

If your server farm has been plagued by space and management problems, blades may help you cut through them. I'm not talking about sharp weapons—a blade is the industry term for a server that fits on a single circuit board, including CPU, memory, and perhaps a local hard disk. Multiple blades are plugged into a chassis, where each blade shares a common power supply, cooling system, and communications backplane. Multiple chassis can then be stacked into racks.

By comparison, the conventional approach for rack-mounted servers involves only one server per chassis. A chassis cannot be smaller than one vertical rack unit (1U, or about 1.75 inches high). This limits you to 42 to 48 servers in a standard seven-foot rack. A typical blade chassis is much higher

than 1U, but several can still be stacked in a rack, allowing upwards of 300 servers per rack, depending on the vendor and configuration.

This compact design offers compelling advantages to anyone operating a high-density server farm where space is at a premium. Indeed, blades are the "next big thing" in servers, and it's probable that any given administrator will have to decide whether to adopt them in the near future.

Simplified Setup

One of the main advantages of blades is evident from a glance at the back of the rack. "If you have 24 traditional servers with external networking and external storage, you might have six or eight cables coming out of each one", says Susan Davis, vice president of Egenera, a two-year-old server blade maker based in Marlboro, MA. "That means you have 128 cables coming out the back of the rack. By aggregating output to redundant controllers, we reduce that to four to eight cables."

Reduced complexity leads to other advantages, Davis notes that ordering, delivering, installing, and configuring a conventional server rack can take 8 to 12 weeks. "Once you've set it up, you don't want to change it. If you have over-provisioned for peak Christmas usage, you have to leave it that way." With blades, you can change from one computing resource to another, remapping a blades' IP addresses and disk resources, and bringing up new applications.

Expanding the server farm is easy as well. Adding or replacing a server simply involves sliding out the chassis and plugging in the new card. Because the subsequent software can be configured remotely, expensive technicians are no longer required on-site. After all, even low-level techs can plug in a board.
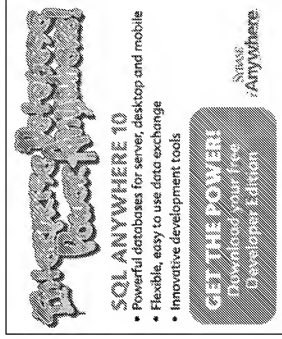
Cost of Ownership

Dick Slansky, analyst at the ARC Advisory Group, a market research and consulting firm in Dedham, MA, studied the total cost of ownership for a blade rack containing 48 blade servers. He found that the blade rack cost 17 to 20 percent less than a traditional rack installation of 48 servers. While the up-front cost might actually be slightly greater, Slansky cites statistics suggesting that, over three years, the blades would save their owners significant amounts of money because they require less floor space and electricity.

Blades reduce labor costs as well. According to many analysts, the standard ratio of administrators to conventional servers has been one to 40. That figure held true even as the cost of servers fell. Blades, however, cost one-seventh as much to administer, Slansky says, smashing the old ratio.

"When you confine [your servers] to a much smaller area, it's easier to see a light blinking on a particular blade, and if a blade goes out you can swap it right out," says Bob Sutherland, an analyst at Technology Business Research, a Hampton, NH-based market research firm.

"It's enterprise users who will push blades," says Slansky. "What made server farms mushroom was the capacity [enterprises] needed to run Internet-based businesses, Web servers make up a big chunk of any server farm." But telecom companies, the engineering community, and certain government agencies will still want processor-intensive servers, mainframes, and supercomputers, Slansky notes. Blades aren't for everyone.

Volume Over Voltage

Blade servers have a performance disadvantage when compared to high-end servers and mainframes, due to the fact that they often employ low-power CPUs. Many of the first blades on the market, such as those from RLX Technologies, used Transmeta's low-wattage Crusoe processor chips. Low-wattage Pentiums have since entered the market as well.

The reason these low-power chips are used has little to do with actual power consumption. Any service provider's server farm will have an adequate power supply or the company won't be in business long. "We haven't focused on lower power," Egenera's Davis says. "Saving a thousand dollars a year on power is not an issue...The idea is to reduce the total cost of ownership by moving from a proprietary Unix architecture to a commodity architecture."

What is an issue for blades, however, is heat dissipation. Today's high-end processors often produce far too much heat for an enclosed, high-density chassis design, leaving low-power chips like those used in laptop computers the only viable option for server blades. Unfortunately, these processors don't deliver the performance of their juice-guzzling cousins, and are typically untried in a server environment.

Ashley Eikenberry, manager for blade marketing at Sun Microsystems, says we have only see the first wave of blade servers. "[Today's models] are single-processor blades focused on the Web tier and the edge of the

network," she says. "Wave two will involve larger blades, with higher-powered processors, and different power and cooling requirements. That will be in the 2004 time frame." Eikenberry says Sun will have a blade product in the second half of 2002. Meanwhile, Egenera claims to be the only blade vendor currently offering two and four-way high performance blade servers.

Searching for Standards

Living up to promises of easy manageability has also been a challenge for first-generation blade servers. "The management software is not mature," notes Gordon Haff, analyst at Illuminata, a data center industry market research firm in Nashua, NH. "That situation, however, is not unique to blades. The ultimate goal is to make a collection of small systems behave in many ways like a single large system could behave, scheduling jobs and adjusting itself automatically in response to changes in the workload. That's not done yet—although the situation is the same with rack-mounted servers."

The biggest disadvantage of blades, however, remains that they involve new hardware that demands new investments. For the purchaser, that means new risks. Lack of hardware standardization means that it can be difficult, if not impossible, to compare equivalent products from multiple vendors. "Any 1U server can [physically] replace any other 1U server," Sutherland explains, "but nothing like that has happened with blades yet."

Standardization is among the top issues vendors are now trying to address. "Customers don't want to invest in a lot of different form factors, so what we've done is built our design around an industry standard and leveraged the momentum of Compact PCI," says Kate O'Neill, blade server product manager at Hewlett Packard. "It won't be just us who supplies blades, but third parties as well. We'll be able to build a catalog of offerings."

But other vendors see HP's solution as little more than a stopgap measure. "Compact PCI is really a telecom standard," says Davis. "We don't see that it provides the capability needed in a data center. HP has chosen that route because they're gunning for Sun, and Sun has owned the [telecom] central office market for years. But no other vendors are jumping on the bandwagon."

One emerging technology that may offer a viable solution is InfiniBand, Davis says.

InfiniBand is a channel-based, switched fabric, point-to-point, interconnect architecture designed especially for servers. It will offer link rates of up to

30Gb. (For more information, see this month's "Access" interview with Dr. Thomas Bradlich of the InfiniBand Trade Association.)

Not everyone is convinced, however. "There is a debate about whether the future backplane will be InfiniBand," says HP's O'Neill. Eikenberry would not comment on what standard Sun would use, except to say that the product would incorporate a Sun processor and operating system.

Making the Call

Compaq's Stevens cautions that whatever standard does finally emerge, it likely won't support any legacy cards you were previously using, such as SSL accelerators. "But the target market doesn't seem to care about that, since everything can be done through software," she adds.

Technology Business Research's Sutherland is less generous. "I don't feel that there is a great deal of demand yet," he says. "It's a cart before the horse thing. Sometimes the customers drive the market, and at other times the vendors try to drive the market. And I think a lot of vendors are waiting for the pool to get deep enough. They learned from the dot-com area to wait until money could be made."

But Davis chalks this up to simple shortsightedness. "Blades aren't just about packing processing into smaller spaces," she says. "Blades are the enabling technology for the next generation of data center computing. The vision involves the disaggregation of resources, the tearing apart of storage, networking, and processing so that they become virtual resources. These resources can come together at points in time to service applications, and then disaggregate to serve other needs."

Davis's vision of high-density computing is an ambitious one, and realistically it's still a few years off. Whether server blades offer advantages for your data center today depends on several factors. But regardless of how they see the standardization issue evolving, all sources agree that blades represent the future for servers—and perhaps even for enterprise computing in general.

Lamont is a freelance writer who lives in San Antonio, Texas. He has been covering the high-tech industry for the last two decades. You can contact him at lwood@texas.net.